END
DATE
FILMED

DTIC

MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

ESTIMATION FOR DIFFUSION PROCESSES
UNDER MISSPECIFIED MODELS

by

Ian W. McKeague

March, 1983

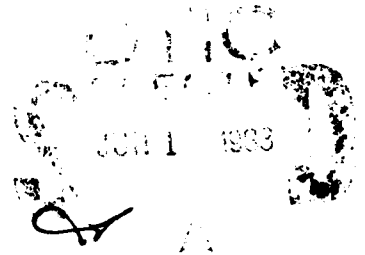The Florida State University
Department of Statistics
Tallahassee, Florida   32306

ESTIMATION FOR DIFFUSION PROCESSES
UNDER MISSPECIFIED MODELS

by

Ian W. McKeague
Florida State University

ABSTRACT

The asymptotic behavior of the maximum likelihood estimator of a para-
meter in the drift term of a stationary ergodic diffusion process is studied
under conditions in which the true drift function and true noise function
do not coincide with those specified by the parametric model.

83 06 01 01

1. Introduction.

Consider the problem of estimating the drift function b(x) of a stationary diffusion process (X_t) given by

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad t \geq 0,$$

where the process is observed over $[0, T]$. The method of maximum likelihood can be used if b(x) is assumed to have a parametric form $f(x, \theta)$, $\theta \in \Theta$. Brown and Hewitt (1975), Kutoyants (1977), Lanska (1979) and Prakasa Rao and Rubin (1981) have shown that the maximum likelihood estimator of θ is consistent and asymptotically normal. Nonparametric methods of estimating b have been developed by Banon (1978) and Geman (1980).

Suppose that a parametric model for the process (X_t) is given by

$$dX_t = f(X_t, \theta)dt + \gamma(X_t)dW_t, \quad t \geq 0.$$

This paper studies the asymptotic behavior of the maximum likelihood estimator of θ under departures of the true drift function b(x) or true noise function $\sigma(x)$ from those specified by the parametric model.

The need for such analysis stems from the desirability of using estimators that are robust under small departures from the underlying model. This kind of analysis is familiar in other settings, for example Huber (1967), White (1981) and Berger and Langberg (1981).

In Section 2 it is shown that the maximum likelihood estimator converges almost surely to a parameter θ^* such that $f(x, \theta^*)$ minimizes an L^2 distance of the parametric family from the true drift function. The L^2 distance is defined with respect to the measure $\gamma^{-2}(x)d\nu(x)$, where ν is the stationary distribution of the process. Asymptotic normality is also established. In Section 3 we discuss a way of estimating the difference in goodness of fit of two separate parametric families to the true drift function. Examples are given at the end of each section.

2. Maximum likelihood estimation under misspecified models.

Let $(X_t, t \geq 0)$ be a stationary, ergodic process which is assumed to be the unique solution of the stochastic differential equation

(2.1)
$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad t \geq 0$$

where X_0 is distributed according to the stationary distribution of the process, b and σ are unknown measurable functions and $(W_t, t \geq 0)$ is a standard Wiener process. Assume that (X_t) has inaccessible boundaries on the state space $(-\infty, \infty)$. Using the notation of Mandl (1968), let

$$B(x) = 2 \int_0^x \frac{b(y)}{\sigma^2(y)}\, dy, \quad p(x) = \int_0^x \exp(-B(y))dy,$$

$$m(x) = 2 \int_0^x \frac{\exp B(y)}{\sigma^2(y)}\, dy,$$

where the integrals are assumed to exist. Provided that $m(+\infty) < \infty$ and $m(-\infty) > -\infty$, the stationary distribution, denoted ν, has distribution function $M^{-1}m(x)$ where $M = m(+\infty) - m(-\infty)$.

Suppose a parametric model is used to estimate the drift function $b(x)$ by the method of maximum likelihood. Let Θ denote a closed bounded interval. A family of measurable drift functions $\{f(x,\theta), \theta \in \Theta\}$ and a measurable noise function $\gamma(x) > 0$ are provided and inference is based on the model

(2.2)
$$dX_t = f(X_t, \theta)dt + \gamma(X_t)dW_t, \quad t \geq 0.$$

The process (X_t) is observed over $[0, T]$. Let μ_θ^T and μ^T denote the measures induced on $C[0, T]$ by process satisfying (2.2) and the process

$$dY_t = \gamma(Y_t)dW_t, \quad t \geq 0$$

$$Y_0 = X_0,$$

respectively. Under conditions given in Liptser and Shiryayev (1977, Theorem 7.19) it follows that $\mu_\theta^T \ll \mu^T$ for all $\theta \in \Theta$ and the log likelihood function $\ell_T(\theta) = \log[d\mu_\theta^T/d\mu^T](X)$ is given by

$$(2.3) \qquad \ell_T(\theta) = \int_0^T \frac{f(X_t,\theta)}{\gamma^2(X_t)} dX_t - \frac{1}{2} \int_0^T \left[\frac{f(X_t,\theta)}{\gamma(X_t)}\right]^2 dt, \quad a.e. \ (\mu^T).$$

A maximum likelihood estimator calculated from $\ell_T(\theta)$ is denoted $\hat\theta_T$.

Assume that $E\left[\dfrac{b(X_0)-f(X_0,\theta)}{\gamma(X_0)}\right]^2 < \infty$ for all $\theta \in \Theta$. and, as a function of

θ, has a unique minimum at $\theta^* \in \Theta$. The following results describe the asymptotic behavior of $\hat\theta_T$ when the observed process satisfies (2.1). The conditions are stated later. g', g'' denote first and second partial derivatives of a function $g(x, \theta)$ with respect to θ.

Theorem 2.1. Under conditions (C1)-(C3), $\hat\theta_T \to \theta^*$ a.s. as $T \to \infty$.

Theorem 2.2. Under conditions (C1)-(C7), $T^{\frac{1}{2}}(\hat\theta_T-\theta^*) \overset{D}{\to} N(0, \Sigma)$ where

$$(2.4) \qquad \Sigma = \frac{2M \int_{-\infty}^\infty g'(y,\theta^*) \int_y^0 \int_{-\infty}^s g'(z,\theta^*)dm(z)dp(s)dm(y)}{\left\{\int_{-\infty}^\infty g''(x,\theta^*)dm(x)\right\}^2},$$

$$(2.5) \qquad g(x, \theta) = \left[\frac{f(x,\theta)}{\gamma(x)}\right]^2 + \sigma^2(x) \frac{\partial}{\partial x}\left(\frac{f(x,\theta)}{\gamma^2(x)}\right).$$

Theorem 2.3. In the special case that the drift function has been correctly specified, i.e. $b(x) = f(x, \theta_0)$ for some $\theta_0 \in \Theta$, then (whether or not the noise function has been correctly specified)

(i) $\hat\theta_T \to \theta_0$ a.s. as $T \to \infty$ under conditions (C1)-(C3);

(ii) $T^{\frac{1}{2}}(\hat{\theta}_T - \theta_0) \overset{\mathcal{D}}{\to} N(0, V)$ under conditions (Cl)-(C9), where

(2.6)
$$V = \frac{\int_{-\infty}^{\infty} \left[\frac{f'(y, \theta_0)}{\gamma(y)}\right]^2 \left[\frac{\sigma(y)}{\gamma(y)}\right]^2 d\nu(y)}{\left\{\int_{-\infty}^{\infty} \left[\frac{f'(x, \theta_0)}{\gamma(x)}\right]^2 d\nu(x)\right\}^2} .$$

Remarks.

(a) The parameter θ^* can often be determined from the moments of the stationary distribution. In general it is difficult to evaluate Σ unless the drift function has been correctly specified (as in Theorem 2.3). Some examples in which θ^* and Σ are explicitly calculated are given at the end of this section.

(b) When both drift and noise functions have been correctly specified V reduces to the previously known formula

$$V_0 = \left(\int_{-\infty}^{\infty} \left[\frac{f'(x, \theta_0)}{\sigma(x)}\right]^2 d\nu_0(x)\right)^{-1} .$$

The Cauchy-Schwarz inequality yields $V \geq V_0$.

(c) Theorem 2.3 (i) can be interpreted as a robustness result for the maximum likelihood estimator; $\hat{\theta}_T$ remains consistent under misspecifications of the noise function.

(d) Lanska (1979) introduced a minimum contrast estimator $\tilde{\theta}_T$ which minimizes

$$\int_0^T h(X_t, \theta) dt$$

where

$$h(x, \theta) = \left[\frac{f(x,\theta)}{\gamma(x)}\right]^2 + \gamma^2(x) \frac{\partial}{\partial x}\left(\frac{f(x,\theta)}{\gamma^2(x)}\right).$$

If h satisfies the conditions (C1)-(C7) in place of g and the function $Eh(X_0, \theta)$ has a unique minimum at θ_1, then using similar proofs to Theorems 2.1 and 2.2, $\tilde{\theta}_T \to \theta_1$ a.s. and $T^{\frac{1}{2}}(\tilde{\theta}_T - \theta_1) \overset{D}{\to} N(0, \Sigma_1)$, where Σ_1 is given by (2.4) with g replaced by h. Even when the drift function is correctly specified, it is possible for $\theta_1 \neq \theta_0$ when the noise function is misspecified, so that $\tilde{\theta}_T$ can fail to be a consistent estimator of θ_0 while $\hat{\theta}_T$ remains consistent.

Conditions.

(C1) $|f(x,\theta_1)-f(x,\theta_2)| \leq J(x)\psi(\theta_1-\theta_2)$, $x \in R$, θ_1, $\theta_2 \in \Theta$, where $E\left[\frac{J(X_0)}{\gamma(X_0)}\right]^2 < \infty$ and $\lim_{\alpha \to 0} \psi(\alpha) = 0$.

(C2) $E\left[\frac{\sigma(X_0)}{\gamma(X_0)}\right]^2 < \infty$ and $|f(x,\theta)| \leq K(x)$, $x \in R$, $\theta \in \Theta$, where $E\left[\frac{K(X_0)}{\gamma(X_0)}\right]^2 < \infty$.

(C3) $f(x, \theta)$ is continuous in (x, θ) and differentiable with respect to θ. There exists $\alpha > 0$ such that $|f'(x,\theta_1)-f'(x,\theta_2)| \leq c(x)|\theta_1-\theta_2|^\alpha$, $x \in R$, θ_1, $\theta_2 \in \Theta$, where $E\left[\frac{\sigma(X_0)c(X_0)}{\gamma^2(X_0)}\right]^2 < \infty$.

(C4) $f(x, \theta)\gamma^{-2}(x)$ has a continuous first partial derivative with respect to x for each $\theta \in \Theta$.

(C5) $\lim_{x \to \pm\infty} f(x, \theta)\gamma^{-2}(x)\exp B(x) = 0$, $\forall$ $\theta \in \Theta$.

(C6) The partial derivatives g', g'', G', G'' exist and are continuous in (x, θ), where $G(x, \theta)$ is defined by

$$(2.7) \qquad G(x, \theta) = \int_0^x \frac{f(y,\theta)}{\gamma^2(y)} \, dy.$$

(C7) $|g''(x,\theta_1)-g''(x,\theta_2)| \leq U(x)\phi(\theta_1-\theta_2)$, $x \in R$, θ_1, $\theta_2 \in \Theta$, where $E \, U(X_0) < \infty$ and $\lim_{\alpha \to 0} \phi(\alpha) = 0$.

(C8) $\lim_{x \to \pm\infty} f'(x,\theta_0)\gamma^{-2}(x)[\exp B(x)] \int_0^x f'(s,\theta_0)\gamma^{-2}(s)ds = 0$.

(C9) $\lim_{x \to \pm\infty} f''(x,\theta_0)\gamma^{-2}(x)\exp B(x) = 0$.

(C10) Condition (C7) with g' in place of g''.

<u>Proof of Theorem 2.1.</u> The proof of strong consistency of $\hat{\theta}_T$ in the correctly specified case given by Prakasa Rao and Rubin (1981, Theorem 4.1) needs only minor modifications to show that $\hat{\theta}_T \to \theta^*$ a.s. as $T \to \infty$ in the misspecified case. The details are given here for completeness. From (2.1) and (2.3) $\ell_T(\theta)$ can be written

$$(2.6) \quad \ell_T(\theta) = -\tfrac{1}{2}\int_0^T \left[\frac{f(X_t,\theta)-b(X_t)}{\gamma(X_t)}\right]^2 dt + \tfrac{1}{2}\int_0^T \left[\frac{b(X_t)}{\gamma(X_t)}\right]^2 dt + \int_0^T \frac{f(X_t,\theta)\sigma(X_t)}{\gamma^2(X_t)} \, dW_t.$$

Denote $I_T(\theta) = \int_0^T \left[\frac{f(X_t,\theta)-b(X_t)}{\gamma(X_t)}\right]^2 dt$. For θ_1, $\theta_2 \in \Theta$,

$$|I_T(\theta_1)-I_T(\theta_2)| \leq \int_0^T \frac{|f(X_t,\theta_1)-f(X_t,\theta_2)|(|f(X_t,\theta_1)|+|f(X_t,\theta_2)|)}{\gamma^2(X_t)} \, dt$$

$$+ 2 \int_0^T \frac{|b(X_t)||f(X_t,\theta_1)-f(X_t,\theta_2)|}{\gamma^2(X_t)}$$

$$\leq 2\psi(\theta_1-\theta_2) \int_0^T \frac{J(X_t)(K(X_t)+b(X_t))}{\gamma^2(X_t)} \, dt,$$

using conditions (C1) and (C2). By the ergodic theorem

$$\frac{1}{T} \int_0^T \frac{J(X_t)(K(X_t)+b(X_t))}{\gamma^2(X_t)} \, dt \to E\left[\frac{J(X_0)(K(X_0)+b(X_0))}{\gamma^2(X_0)}\right] \quad \text{a.s.}$$

as $T \to \infty$, using conditions (C1) and (C2) again. Thus, there exists a r.v. C^*, which is finite a.s. and does not depend on θ, such that

$$|I_T(\theta_1)-I_T(\theta_2)| \le C^* T \psi(\theta_1-\theta_2), \quad \text{for all } \theta_1, \theta_2 \in \Theta, \quad T \ge 0.$$

Similarly, using condition (C2) again, it follows that there exists a r.v. D^*, which is finite a.s. and does not depend on θ, such that $I_T(\theta) \le D^* T$, for all $\theta \in \Theta$, $T \ge 0$. Thus, $\{\frac{1}{T} I_T(\cdot), T \ge 0\}$ is equicontinuous and uniformly bounded a.s. as a family of functions of θ. By the Arzela-Ascoli theorem this family is relatively compact (a.s.) in the space of continuous functions on Θ provided with the supremum norm. Therefore, by the ergodic theorem

$$\frac{1}{T} I_T(\theta) \xrightarrow{\text{a.s.}} E\left[\frac{f(X_0,\theta)-b(X_0)}{\gamma(X_0)}\right]^2,$$

uniformly in $\theta \in \Theta$ as $T \to \infty$.

Now consider the second term in (2.6). By condition (C2) and the ergodic theorem

$$\frac{1}{T} \int_0^T \left[\frac{b(X_t)}{\gamma(X_t)}\right]^2 dt \xrightarrow{\text{a.s.}} E\left[\frac{b(X_0)}{\gamma(X_0)}\right]^2 \quad \text{as } T \to \infty.$$

Next, using Lemma 4.3 of Prakasa Rao and Rubin (1981) it follows that, under conditions (C1)-(C3),

$$\frac{1}{T} \int_0^T \frac{f(X_t,\theta)\sigma(X_t)}{\gamma^2(X_t)} \, dW_t \xrightarrow{\text{a.s.}} 0, \quad \text{uniformly in } \theta \in \Theta \text{ as } T \to \infty.$$

Thus

$$(2.7) \qquad \frac{1}{T} \, \ell_T(\theta) \xrightarrow{a.s.} -\tfrac{1}{2} E\left[\frac{f(X_0,\theta)-b(X_0)}{\gamma(X_0)}\right]^2 + \tfrac{1}{2} E\left[\frac{b(X_0)}{\gamma(X_0)}\right]^2 ,$$

uniformly in $\theta \in \Theta$ as $T \to \infty$. Since the r.h.s. of (2.7) has a unique maximum at $\theta^* \in \Theta$ and $\hat{\theta}_T$ maximizes $\frac{1}{T} \, \ell_T(\theta)$, it is easily proved that $\hat{\theta}_T \to \theta^*$ a.s. as $T \to \infty$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ $\square$

<u>Proof of Theorem 2.2.</u> The approach used by Prakasa Rao and Rubin (1981) to find the asymptotic distribution of $\hat{\theta}_T$ in the correctly specified case does not extend to the misspecified case. Rather, the proof of this theorem uses the technique, introduced by Lanska (1979), of expressing $\ell_T(\theta)$ in terms of Lebesgue integrals.

The function $G(x, \theta)$ defined in (2.7) has a continuous second partial derivative with respect to x for each $\theta \in \Theta$ by condition (C4). Applying Itô's formula, it follows that

$$G(X_T, \theta) = G(X_0, \theta) + \int_0^T \left[\frac{b(X_t)f(X_t,\theta)}{\gamma^2(X_t)} + \tfrac{1}{2}\sigma^2(X_t)G'(X_t,\theta)\right]dt$$

$$+ \int_0^T \frac{\sigma(X_t)f(X_t,\theta)}{\gamma^2(X_t)} \, dW_t .$$

Then, using (2.1) and (2.3),

$$\ell_T(\theta) = \int_0^T \frac{f(X_t,\theta)b(X_t)}{\gamma^2(X_t)} \, dt + \int_0^T \frac{f(X_t,\theta)\sigma(X_t)}{\gamma^2(X_t)} \, dW_t - \tfrac{1}{2}\int_0^T \left[\frac{f(X_t,\theta)}{\gamma(X_t)}\right]^2 dt$$

$$(2.8)$$

$$= G(X_T,\theta) - G(X_0,\theta) - \tfrac{1}{2}\int_0^T g(X_t,\theta)dt ,$$

where g is defined in (2.5). Expand $\ell_T^!(\theta)$ about $\hat{\theta}_T$,

$$\ell_T^!(\theta^*) = \ell_T^!(\hat{\theta}_T) + (\theta^*-\hat{\theta}_T)\ell_T^{!!}(\tilde{\theta}_T), \quad \text{where } |\tilde{\theta}_T-\theta^*| \leq |\hat{\theta}_T-\theta^*|.$$

Consider

$$T^{-\frac{1}{2}}\ell_T^!(\theta^*) = T^{-\frac{1}{2}}(G'(X_T,\theta^*) - G'(X_0,\theta^*)) + T^{-\frac{1}{2}} \int_0^T g'(X_t,\theta^*)dt.$$

Using the stationarity of (X_t),

$$T^{-\frac{1}{2}}(G'(X_T,\theta^*) - G'(X_0,\theta^*)) \xrightarrow{p} 0 \quad \text{as } T \to \infty.$$

Using integration by parts and condition (C5) it can be shown that

$$(2.9) \qquad E\, g(X_0, \theta) = E\left[\frac{f(X_0,\theta)-b(X_0)}{\gamma(X_0)}\right]^2 - E\left[\frac{b(X_0)}{\gamma(X_0)}\right]^2,$$

and since the right hand side of this expression is minimized at $\theta^* \in \Theta$, it follows that $E\, g(X_0, \theta)$ is minimized at $\theta^* \in \Theta$ and $E\, g'(X_0, \theta^*) = 0$. Then, by Mandl (1968, p. 94)

$$T^{-\frac{1}{2}} \int_0^T g'(X_t, \theta^*) \xrightarrow{D} N(0, \Delta),$$

where

$$\Delta = \frac{2}{M} \int_{-\infty}^{\infty} g'(y,\theta^*) \int_y^0 \int_{-\infty}^s g'(z,\theta^*)dm(z)dp(s)dm(y).$$

Thus $T^{-\frac{1}{2}}\ell_T^!(\theta^*) \xrightarrow{D} N(0, \Delta)$. By the ergodic theorem and condition (C6)

$$\frac{1}{T} \ell_T^{!!}(\theta^*) \xrightarrow{p} E\, g''(X_0, \theta^*).$$

Next, using conditions (C6), (C7) and the fact that $\tilde{\theta}_T \to \theta^*$ a.s.

$$\frac{1}{T}(\ell_T''(\tilde{\theta}_T) - \ell_T''(\theta^*)) \overset{P}{\to} 0, \quad \text{as } T \to \infty.$$

Thus

$$\frac{1}{T} \ell_T''(\tilde{\theta}_T) \overset{P}{\to} E\ g''(X_0, \theta^*) = \frac{1}{M} \int_{-\infty}^{\infty} g''(x,\theta^*)dm(x),$$

as $T \to \infty$. We conclude that

$$T^{\frac{1}{2}}(\hat{\theta}_T - \theta^*) \overset{D}{\to} N(0, \Sigma),$$

where Σ is given in (2.4). $\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad$ $\square$

<u>Proof of Theorem 2.3.</u> Suppose that $b(x) = f(x, \theta_0)$, where $\theta_0 \in \Theta$. Then $\theta^* = \theta_0$ and (a) follows directly from Theorem 2.1. The proof of (b) consists in showing that Σ in (2.4) reduces to V given in (2.6). Note that

$$g'(x, \theta_0) = 2b(x)f'(x,\theta_0)\gamma^{-2}(x) + \sigma^2(x) \frac{\partial}{\partial x}[f'(x,\theta_0)\gamma^{-2}(x)].$$

Using integration by parts,

$$\int_{-\infty}^{s} \sigma^2(z) \frac{\partial}{\partial z} [f'(z,\theta_0)\gamma^{-2}(z)]dm(z)$$

$$= 2f'(s,\theta_0)\gamma^{-2}(s)\exp B(s) - 2 \int_{-\infty}^{s} b(z)f'(z,\theta_0)\gamma^{-2}(z)dm(z),$$

so that

$$\int_{-\infty}^{s} g'(z,\theta_0)dm(z) = 2f'(s,\theta_0)\gamma^{-2}(s)\exp B(s),$$

and

$$\int_y^0 \int_{-\infty}^s g'(z,\theta_0)dm(z)dp(s) = 2 \int_y^0 f'(s,\theta_0)\gamma^{-2}(s)ds.$$

Using integration by parts with condition (C8),

$$\int_{-\infty}^\infty \sigma^2(y) \frac{\partial}{\partial y}[f'(y,\theta_0)\gamma^{-2}(y)] \int_y^0 f'(s,\theta_0)\gamma^{-2}(s)ds \, dm(y)$$

$$= -2\int_{-\infty}^\infty f'(y,\theta_0)\gamma^{-2}(y) \frac{\partial}{\partial y}\left[\int_y^0 f'(s,\theta_0)\gamma^{-2}(s)ds \, \exp B(y)\right]dy$$

$$= \int_{-\infty}^\infty \left[\frac{f'(y,\theta_0)}{\gamma(y)}\right]^2 \left[\frac{\sigma(y)}{\gamma(y)}\right]^2 dm(y) - 2\int_{-\infty}^\infty b(y)f'(y,\theta_0)\gamma^{-2}(y) \int_y^0 f'(s,\theta_0)\gamma^{-2}(s)ds \, dm(y).$$

It follows that the numerator of Σ reduces to

$$4M \int_{-\infty}^\infty \left[\frac{f'(y,\theta_0)}{\gamma(y)}\right]^2 \left[\frac{\sigma(y)}{\gamma(y)}\right]^2 dm(y).$$

Now consider the denominator of Σ.

$$g''(x, \theta_0) = 2\left[\frac{f'(x,\theta_0)}{\gamma(x)}\right]^2 + 2f(x,\theta_0)f''(x,\theta_0)\gamma^{-2}(x) + \sigma^2(x) \frac{\partial}{\partial x}[f''(x,\theta_0)\gamma^{-2}(x)].$$

Using integration by parts and condition (C9)

$$\int_{-\infty}^\infty \sigma^2(x) \frac{\partial}{\partial x}[f''(x,\theta_0)\gamma^{-2}(x)]dm(x) = -2 \int_{-\infty}^\infty f(x,\theta_0)f''(x,\theta_0)\gamma^{-2}(x)dm(x)$$

and it follows that

$$\int_{-\infty}^\infty g''(x,\theta_0)dm(x) = 2 \int_{-\infty}^\infty \left[\frac{f'(x,\theta_0)}{\gamma(x)}\right]^2 dm(x).$$

This completes the proof of the theorem. □

Examples.

1. This example presents a misspecified drift function but correctly specified noise function. Suppose that the observed process satisfies

$$dX_t = -X_t dt + \sqrt{2}dW_t,$$

where X_0 has a N(0, 1) distribution, the stationary distribution of (X_t). Estimates are calculated from the parametric model

$$dX_t = -\theta X_t^3 dt + \sqrt{2}dW_t.$$

The parameter θ^* which minimizes $E(X_0 - \theta X_0^3)^2$ is given by

$$\theta^* = \frac{EX_0^4}{EX_0^6} = \frac{1}{5}.$$

By Theorem 2.1, $\hat{\theta}_T \to \frac{1}{5}$ a.s. as $T \to \infty$. We also have $g(x, \theta) = \frac{1}{2}\theta^2 x^6 - 3\theta x^2$, so that $g'(x, \theta^*) = \frac{1}{5} x^6 - 3x^2$. Using repeated integration by parts

$$\int_{-\infty}^{s} g'(z,\theta^*)dm(z) = \int_{-\infty}^{s} (\frac{1}{5} z^6 - 3z^2)e^{-z^2/2}dz = -(\frac{1}{5} s^5 + s^3)e^{-s^2/2}$$

so that

$$\int_{y}^{0} \int_{-\infty}^{s} g'(z,\theta^*)dm(z)dp(s) = \frac{1}{30} y^6 + \frac{1}{4} y^4$$

and

$$\int_{-\infty}^{\infty} g'(y,\theta^*) \int_{y}^{0} \int_{-\infty}^{s} g'(z,\theta^*)dm(z)dp(s)d\nu(y) = \int_{-\infty}^{\infty} (\frac{1}{5} y^6 - 3y^2)(\frac{1}{30} y^6 + \frac{1}{4} y^4)d\nu(y)$$

$$= \frac{1}{150} EX_0^{12} + \frac{1}{20} EX_0^{10} - \frac{1}{10} EX_0^8 - \frac{3}{4} EX_0^6$$

$$= 94.8.$$

Also

$$\int_{-\infty}^{\infty} g''(x,\theta^*)d\nu(x) = EX_0^6 = 15.$$

Thus $\Sigma = \dfrac{(2)(94.8)}{(15)^2} = .84$, and by Theorem 2.2 we have $T^{\frac{1}{2}}(\hat{\theta}_T - \frac{1}{5}) \overset{D}{\to} N(0, .84)$.
The asymptotic variance of $\hat{\theta}_T$ is less than in the correctly specified case
for which $T^{\frac{1}{2}}(\hat{\theta}_T - 1) \overset{D}{\to} N(0, 1)$.

2. Our second example has a correctly specified drift function and
a misspecified noise function. The observed process is the same as in the
first example but the parametric model is given by

$$dX_t = -\theta X_t dt + \left(\frac{2}{1+X_t^2}\right)^{\frac{1}{2}} dW_t.$$

Theorem 2.3 yields $T^{\frac{1}{2}}(\hat{\theta}_T - 1) \overset{D}{\to} N(0, 2.75)$. *The asymptotic variance of* $\hat{\theta}_T$
has almost tripled due to the misspecified noise.

3. Discriminating between separate families of drift functions.

Let (X_t) satisfy (2.1) and assume throughout this section $\sigma(x) \equiv 1$.
Suppose that two parametric models for this process have been suggested.
It is required to decide in favor of the model which best fits the observed
trajectory $\{X_t, 0 \le t \le T\}$.

Let $\{f_1(x,\theta): \theta \in \Theta\}$, $\{f_2(x,\phi): \phi \in \Phi\}$ be distinct families of drift
functions, where Θ, Φ are closed bounded intervals. A reasonable way to
compare the goodness of fit of these families to the true drift function
$b(x)$ is to estimate the parameter

$$\Delta = E[f_2(X_0, \phi^*) - b(X_0)]^2 - E[f_1(X_0, \theta^*) - b(X_0)]^2.$$

In this section we introduce an estimator for Δ. The noise function is assumed to be correctly specified; that is $\gamma(x) \equiv 1$.

Let $\ell_T^{(1)}(\theta)$, $\ell_T^{(2)}(\phi)$ denote the log likelihoods for the two models. Define

$$\hat{\Delta}_T = \frac{2}{T}[\ell_T^{(1)}(\hat{\theta}_T) - \ell_T^{(2)}(\hat{\phi}_T)].$$

Given that f_1, f_2 satisfy conditions (C1)-(C3), it follows from (2.7) in the proof of Theorem 2.1 that $\hat{\Delta}_T \to \Delta$ a.s. as $T \to \infty$. The following result shows that $\hat{\Delta}_T$ is asymptotically normal. Conditions from section 2 are used interchangeably between the two families of drift functions indexed by θ and ϕ.

<u>Theorem 3.1.</u> Suppose that f_1, f_2 satisfy conditions (C1)-(C6) and (C10), where g_1, g_2 are given by (2.5) with $f = f_1$, f_2 respectively, $\gamma(x) \equiv 1$. Then $T^{\frac{1}{2}}(\hat{\Delta}_T - \Delta) \xrightarrow{D} N(0, \Sigma_2)$, where

$$\Sigma_2 = \frac{2}{M} \int_{-\infty}^{\infty} [g_2(y, \phi^*) - g_1(y, \theta^*) - \Delta] \int_{y}^{0} \int_{-\infty}^{s} [g_2(z, \phi^*) - g_1(z, \theta^*) - \Delta] dm(z) dp(s) dm(y).$$

<u>Proof.</u> From (2.8)

(3.1)
$$T^{\frac{1}{2}}(\hat{\Delta}_T - \Delta) = T^{-\frac{1}{2}} \int_0^T [g_2(X_t, \hat{\phi}_T) - g_1(X_t, \hat{\theta}_T) - \Delta] dt$$
$$+ 2T^{-\frac{1}{2}}[G_1(X_T, \hat{\theta}_T) - G_1(X_0, \hat{\theta}_T) - G_2(X_T, \hat{\phi}_T) + G_2(X_0, \hat{\phi}_T)].$$

The second term on the right hand side of (3.1) converges to 0 in probability. The first term is written

$$T^{-\frac{1}{2}} \int\limits_0^T [g_2(X_t,\phi^*) - g_1(X_t,\theta^*) - \Delta]dt$$

$$+ T^{-\frac{1}{2}} \int\limits_0^T [g_2(X_t,\hat{\phi}_T) - g_2(X_t,\phi^*)]dt$$

$$+ T^{-\frac{1}{2}} \int\limits_0^T [g_1(X_t,\theta^*) - g_1(X_t,\hat{\theta}_T)]dt$$

$$= A_T + B_T + C_T.$$

From (2.9),

$$E[g_2(X_0,\phi^*) - g_1(X_0,\theta^*) - \Delta] = 0$$

so, by Mandl (1968, p. 94), $A_T \overset{D}{\to} N(0, \Sigma_2)$. Next, consider C_T. Expanding g_1 in a neighborhood around θ^*,

$$g_1(x,\theta) = g_1(x,\theta^*) + (\theta-\theta^*)g_1'(x,\tilde{\theta}),$$

where $|\tilde{\theta}-\theta^*| \le |\theta-\theta^*|$. This gives

$$C_T = T^{\frac{1}{2}}(\hat{\theta}_T-\theta^*) \cdot \frac{1}{T} \int\limits_0^T g_1'(X_t,\tilde{\theta}_T)dt.$$

But,

$$\frac{1}{T} \int\limits_0^T g_1'(X_t,\tilde{\theta}_T)dt = \frac{1}{T} \int\limits_0^T g_1'(X_t,\theta^*)dt$$

(3.2)

$$+ \frac{1}{T} \int\limits_0^T [g_1'(X_t,\tilde{\theta}_T) - g_1'(X_t,\theta^*)]dt.$$

By the proof of Theorem 2.2, $E\, g_1'(X_0,\theta^*) = 0$, so by the ergodic theorem the first term in (3.2) converges to 0 a.s. as $T \to \infty$. From condition (C10),

the second term in (3.2) is bounded above by

$$\phi(\bar{\theta}_T - \theta^*) \cdot \frac{1}{T} \int_0^T U(X_t) dt,$$

which converges to 0 a.s. since $|\bar{\theta}_T - \theta^*| \leq |\hat{\theta}_T - \theta^*|$, $\hat{\theta}_T \to \theta^*$ a.s. and

$\frac{1}{T} \int_0^T U(X_t) dt \xrightarrow{a.s.} EU(X_0)$ as $T \to \infty$. Thus $C_T \to 0$ a.s. and similarly $B_T \to 0$ a.s..

This completes the proof of the theorem. $\square$

Example.

Consider the two models

(3.3) $$dX_t = -\theta X_t dt + dW_t,$$

(3.4) $$dX_t = -\phi X_t^3 dt + dW_t,$$

and suppose that the observed process satisfies (3.3) with $\theta = \theta_0 > 0$. Some involved but routine calculations give that $\Delta = .2\theta_0$ and $\Sigma_2 = .64\theta_0 + 9.36\theta_0^{-3}$. Note that $\Sigma_2 \to \infty$ as $\theta_0 \to 0$. The poor performance of $\hat{\Delta}_T$ for small θ_0 is to be expected since, as $\theta_0 \to 0$, the drift function has less effect on the dynamics of the process so it is harder to discriminate between the two models.

Acknowledgement.

REFERENCES

Banon, G. (1978). Nonparametric identification for diffusion processes. Siam J. on Control and Optimization 16, 380-395.

Berger, R.L. and Langberg, N.A. (1981). Linear least squares estimates and nonlinear means. Florida State University Statistics Report M-573.

Brown, B.M. and Hewitt, J.I. (1975). Asymptotic likelihood theory for diffusion processes. J. Appl. Prob. 12, 228-238.

Geman, S. (1980). An application of the method of sieves: functional estimator for the drift of a diffusion. Colloquia Mathematica Societatis János Bolyai 32. Nonparametric Statistical Inference, Budapest (Hungary).

Huber, P.J. (1967). The behavior of maximum likelihood estimates under non-standard conditions. Proc. Fifth Berkeley Symp. Math. Statist. and Prob. 1, 221-233.

Kutoyants, Yu. A. (1977). Estimation of the drift coefficient parameter of a diffusion process in the smooth case. Theory Prob. Appl. 22, 399-406.

Lanska, V. (1979). Minimum contrast estimation in diffusion processes. J. Appl. Prob. 16, 65-75.

Mandl, P. (1968). Analytical Treatment of One-Dimensional Markov Processes. Academia, Prague; Springer-Verlag, Berlin.

Prakasa Rao, B.L.S. and Rubin, H. (1981). Asymptotic theory of estimation in non-linear stochastic differential equations. Sankhyā Ser. A, to appear.

White, H. (1981). Consequences and detection of misspecified nonlinear regression models. J. Amer. Statist. Assoc. 76, 419-433.

Liptser, R.S. and Shiryayev, A.N. (1977). Statistics of Random Processes I, Springer-Verlag, New York.

REPORT DOCUMENTATION PAGE

1. REPORT NUMBER FSU M653 USARO D-59	2. GOVT ACCESSION NO. A128 799	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and subtitle)		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER FSU Statistics Report No. M653
7. AUTHOR(s) Ian W. McKeague		8. CONTRACT OR GRANT NUMBER(s) USARO Grant DAAG29-82-K-0168
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Florida State University Tallahassee, FL 32306		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS U.S. Army Research Office - Durham P.O. Box 12211 Research Triangle Park, NC 27709		12. REPORT DATE March, 1983
		13. NUMBER OF PAGES 17
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this report)

Distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS

Diffusion processes, maximum likelihood estimates, misspecified models.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The asymptotic behavior of the maximum likelihood estimator of a parameter
in the drift term of a stationary ergodic diffusion process is studied under
conditions in which the true drift function and true noise function do not
coincide with those specified by the parametric model.

ME
8